REMARKS

Pending claims 1 to 17 have been canceled and replaced with new claims 18 to 33. No new matter has been added.

In the Official Action, claims 12-17 were objected to under 37 C.F.R. 1.75(c) as allegedly being in improper form as multiple dependent claims that do not refer to other claims in the alternative only. Claims 12-17 have been canceled, thereby obviating this objection.

Claims 1-11 were rejected to under 35 U.S.C. 112, second paragraph, as allegedly being indefinite. The Examiner also noted that claims 12-17 are indefinite and would be rejected under 35 U.S.C. 112, second paragraph, if the improper multiple dependencies were eliminated. Claims 1-17 have been canceled, thereby obviating these rejections.

In addition, claims 1-5 and 7-11 were rejected under 35 U.S.C. 102(e) as allegedly being anticipated by Aihara et al. (US 2003/0065603) ("Aihara") and claim 6 was rejected under 35 U.S.C. 103(a) as allegedly being obvious over the teachings of Aihara. Claims 12-17 have not been treated on their merits in view of the prior art. Claims 1-17 have been canceled and replaced by new claims 18-33, thereby obviating these rejections. New claims 18-33 are believed to comply with the requirements of 35 U.S.C. 112, 35 U.S.C. 102, and 35 U.S.C. 103 and to be in condition for allowance for the reasons given below.

While canceled claims were directed to a controller for controlling a system, new claims 18-33 are directed to a method of controlling a system to optimize an objective function thereof, the system being capable of performing a plurality of candidate actions and of monitoring response performances of a performance of a respective candidate action. The method of the present invention monitors the response performance of a candidate action that is chosen to be performed, stores, according to candidate action performed, a representation of the monitored response performance, and then chooses which of the plurality of candidate actions is next performed so as to optimize the objective function. The choice is made by assessing, using the probability distribution of the response performance of all of the plurality of candidate actions, which candidate action is likely to result in the lowest expected growth in regret after the chosen candidate action is performed. Once this choice has been made, the

method is repeated. The feature of lowest expected growth in regret can be found in original claim 2.

Thus, the chosen candidate action is performed, its response is monitored, and the next choice of candidate action is made, using the probability distribution of the response performance of all of the plurality of candidate actions (including the one just performed), which is again likely to result in the lowest expected growth in regret after this chosen candidate action is performed.

Applicant notes that the term "regret" is used in the claims to describe the shortfall in response performance between always performing the true best candidate action and actually performing the candidate actions chosen to be performed. It is evident from this definition that regret is a cumulative metric. The true best candidate action is not indicative of a true best candidate, it is the true best action that could be taken given everything known about the current circumstances and based on an infinite amount of observational data such that the statistical confidence in the decision process on the true best action is 100% certain that it is actually the best action.

New claims 18-33 include such features and are believed to be definite, novel and nonobvious for the reasons given below.

In the method of new independent claim 18, no prior knowledge about each performance scenario is assumed and no prior knowledge of the environment governing the choice of candidate action is assumed until the moment immediately preceding the performance of that chosen candidate action. There is also no discretion about whether or not a candidate action is going to be performed at all; instead every single available opportunity for candidate action performance can be used by choosing one of the number of available candidate actions. This provides a mathematical freedom by making one of a number of choices right now, for each opportunity as it presents itself. Thus, the choosing is coupled to the performance event in real time (*i.e.*, both the choosing and performance must be taken in the present). The present invention is therefore able to optimally balance investment in new performances versus exploiting the current knowledge about performances observed from historical interaction scenarios. This gives the present invention the ability to make ongoing optimal decisions about investments in learning versus the exploitation of current knowledge.

The claimed method also does not have separate decision and execution phases. The method simultaneously explores and exploits the behavior of the interaction environment in a way that delivers ongoing greatest benefit relative to a perfect decision system at all times during its use. It is therefore a method that is suited to complex interaction environments in which there may be very large numbers of parameters that influence the potential response performances, and for which methods for general mapping the response behavior over the input envelope would take too long, be too expensive, where the response performance is known to change over time in a way that erodes prior learning, where the conditions relating to the next decision scenario are not known until the time the performance takes place, and/or are otherwise impractical. Under these conditions, the method bases each next choice upon the estimated value contributions of exploration and exploitation respectively, for each candidate action, under the prevailing conditions.

By continuously evaluating the uncertainties in its knowledge combined with the apparent potential value delivered by the exploitation of the current knowledge within a unified value metric, the claimed method is designed to improve in choice-making efficiency at the highest possible rate, with each new choice that is made. In this way, the method can operate with optimal performance over any time period, given changes that may be occurring over time, or changes in the available set of candidate actions, such as new candidate actions being introduced or existing candidate actions being excluded, changes to the characteristics of those candidate actions, or changes to any other characteristics of the decision scenarios.

With the present method, the activities of exploration and exploitation can take place simultaneously. Every choice is made on the basis of the value benefit that may be realized from performing the next candidate action. The method explicitly considers the benefit of exploiting what has been seen to work well in the past versus the benefit-risk of ignoring other candidate actions that may in fact prove to be better were more exemplars available for appraisal. 'Regret' is a term that relates the loss of the system which results from non-optimal decisions, where 'non-optimal decisions' are compared to an imaginary system that has access to an infinite number of observations which reflect the current conditions, and which is therefore able to make perfect decisions, with 100% confidence, every time. As such the appraisal of 'regret' in each and every decision directly drives whether the decision is to be biased in favor of exploitation or in favor of further exploration. There is no pre-programmed

activity taking place, no-prior knowledge about the types of future decision scenarios is assumed, and at no time is the future execution of actions based upon a predefined set of decisions which specify the activities for the next period.

Applicant further notes that the claimed method is in complete contrast to the methodology disclosed in Aihara, which relates to a process for modeling the risk in an advertisement purchase when building an advertisement portfolio. The Aihara process involves considering the requirements of a 'sponsor' (who intends to exploit advertising products) and for building an optimal portfolio of advertising products such that the sponsor's success criteria are satisfied, while at the same time minimizing the risk.

Aihara also discloses having a number of discrete types and numbers of advertisement product options. An advertisement portfolio is then selected that comprises a limited subset of those available product options that meet some specific criteria. In selecting a subset of these products from an available set, preconditions are set by a sponsor (or investor). Paragraphs [0076]-[0087] explain the use of 'User Input Purchasing Conditions' which depend upon known characteristics of each future opportunity as a basis for pre-selection and ranking, and explains how these criteria are applied to the known set of advertising opportunities. Thus, the Aihara system allows for the user to specify particular characteristics that will be used as pre-screening acceptance or rejection criteria for any given advertisement product opportunity. Paragraph [0206] explains the concept of the Advertisement Portfolio Model and how it represents a commitment with respect to a subset of opportunities for an available full set.

Applicant notes that Aihara selects the products all at the same time (in a single collective decision event) to exploit their complimentary risk profiles, such that the overall risk exposure is minimized, given the required performance outputs. In particular, the system employs prior knowledge about the types of advertisement product that exist and must have this information in advance in order to define the portfolio. In addition, the system selects a subset of the available opportunities in which to execute an action and it will therefore know that there are a finite number of each type of advertisement product.

Aihara therefore has a transaction decision phase in which the desired outputs (mainly metrics relating to advertising return and risk) are analyzed with respect to the adoption or rejection of specific advertisement product opportunities occurring at known times in the

future. The optimum combination of these products found during this decision phase are then purchased and programmed for execution. The decision freedom being exploited here is represented by a set of buy/no-buy decisions that will be executed over a predefined upcoming period (the set of decisions to be made in a block, all at one time). This is in complete contrast to the present invention where the decision freedom is the choice from all candidate actions of the action to be taken and performed now.

Aihara then proceeds to an exploitation or execution phase in which the specific advertisement products purchased at the end of the transaction decision phase are exploited in a deterministic manner through a programmed roll-out. This execution phase is separated in time from the transaction decision phase.

Thus, there are two quite distinct phases that occur in a particular sequence in Aihara, the transaction decision phase being required to take place before the exploitation phase. Aihara's transaction decision process does not and indeed can not consider the merit of exploration (as opposed to exploitation) and the benefit of improved knowledge as a contributor to the outcome value. In the Aihara system, the estimated or assumed probability distributions are used solely for the purposes of assessing risk and expected return at that time. At no stage in the Aihara process are the probability distributions relating to the outcomes of specific actions considered deliberately for the purposes of understanding how taking those decisions will improve the knowledge state of the system, and improve the systems ability to support future decisions. These characteristics of the claimed method whereby a real-time learning device provides continuous convergence through optimization towards the 'perfect decision system' is not taught by Aihara. On the contrary, the Aihara system simply takes a passive approach to analysis, attempting to efficiently utilize all information existing at that time point, while ignoring how future actions will influence the confidence in future decision-making.

Moreover, the Aihara system is not designed to maximize ongoing learning efficiency. It is instead designed to meet specific expectation return criteria while minimizing risk exposure (either measured as a monetary loss, or loss in expectation advertising return value, over the appraisal period, with a specific confidence threshold). Although Aihara refers to the use of available observational (statistical) data in many sections, this data is not employed in an incremental way, as and when new data becomes

available, such that the system decision performance converges to that of an optimal decision system as rapidly as possible.

The Examiner suggests that Aihara teaches about the 'costs' and 'losses' with respect to presenting specific candidate options. However, Aihara and Applicant are using the terms to mean quite different things. Aihara may discuss the 'costs', 'gains' or 'losses' associated with particular decisions, but these metrics refer to differences between actual observed events and the outcomes predicted by Aihara's risk management system. As such they represent system prediction errors, and are employed to improve the estimates of risk for future predictions. These are not the same as costs or gains that are discussed in the present application that relate to accrued positive and negative estimates of benefit with respect to taking specific candidate courses of action. These metrics are part of the analysis of expectation benefit which is used in appraising the next candidate action. They do not relate to actual prediction errors of the system as in Aihara's system.

In summary, the Aihara system makes no mention of choosing an advertisement or any other product likely to result in the lowest expected growth in regret after the chosen product is performed. Aihara makes no mention of and contains no appreciation of any value in the shortfall in response performance between always performing the true best product and actually performing the product chosen to be performed.

For the above reasons, it is submitted that the invention of new claims 18-33 filed herewith are novel and nonobvious over the disclosure of Aihara. A Notice of Allowability is thus solicited.

Date: July 25, 2005

Vincent J. Roccia
Registration No. 43,887

Woodcock Washburn LLP
One Liberty Place - 46th Floor
Philadelphia PA 19103
Telephone: (215) 568-3100
Facsimile: (215) 568-3439